

KEY_Practice16_Pandas-Subsetting-I

December 10, 2021

1 Practice: Subsetting Pandas DataFrames I

For this practice, let's use the iris dataset:

```
[ ]: # import the pandas package
import pandas as pd
# set the path
path = 'https://raw.githubusercontent.com/GWC-DCMB/curriculum-notebooks/master/'
# this is where the file is located
filename = path + 'SampleData/iris.csv'
# load the iris dataset into a DataFrame
iris = pd.read_csv(filename)
```

Get familiar with the dataset:

```
[ ]: # take a look at the beginning

iris.head()
```

Try subsetting on columns:

```
[ ]: # subset the species column

iris['species']
```

```
[ ]: # subset the sepal_length and sepal_width columns

iris[['sepal_length', 'sepal_width']]
```

```
[ ]: # here's a list of columns we want to keep
columns = ['species', 'petal_width']
# now subset the iris dataframe using the list
iris[columns]
```

Try subsetting on rows:

```
[ ]: # select the first 10 rows
iris.iloc[:10]
# iris.head(10) # alternative method
```

```
[ ]: # think of another way to select the first 10 rows
iris.head(10)
# iris.iloc[:11] # alternative method
```

```
[ ]: # select rows 15 through 20
iris.iloc[15:21]
```

```
[ ]: # select row 30
iris.iloc[30]
```

Bonus: Do you think it's possible to subset with both rows AND columns at the same time? Why or why not?

Feel free to experiment in the space below:

```
[ ]:
```